DOCUMENT RESUME

ED 456 145 TM 033 207

AUTHOR	Madaus, George F.
TITLE	A Brief History of Attempts To Monitor Testing. NBETPP Statements, Volume 2, Number 2.
INSTITUTION	National Board on Educational Testing and Public Policy, Chestnut Hill, MA.
SPONS AGENCY	Ford Foundation, New York, NY.
PUB DATE	2001-02-00
NOTE	17p.
AVAILABLE FROM	For full text: http://www.nbetpp.bc.edu.
PUB TYPE	Opinion Papers (120) -- Reports - Descriptive (141)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	Educational History; Elementary Secondary Education; *Standardized Tests; Supervision; *Test Use; Testing Programs
IDENTIFIERS	*Monitoring

ABSTRACT

This paper provides a brief history of attempts to monitor testing in the United States. It describes proposals for monitoring from the first attempts in the 1920s to similar proposals in the 1990s. The discussion focuses on: (1) Giles Ruche's proposal for a consumer research bureau on tests; (2) Oscar K. Burros' reviews of tests and efforts to establish a more active test monitoring agency; (3) the call of the American Psychological Association for a Bureau of Test Standards and a Seal of Approval; (4) the Project on the Classification of Exceptional Children's recommendation for a National Bureau of Standards for Psychological Tests and Testing; and (5) the efforts of various professional organizations to establish standards for test development and use. The concept of monitoring tests and the impact of testing programs has a long history, but was not translated into practice until the formation in 1998 of the National Board on Educational Testing and Public Policy. The National Board has finally begun the process of independently monitoring tests and testing programs that has been called for since the 1920s. (Contains 28 endnotes.) (SLD)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 456 145

The National Board on Educational Testing and Public Policy

NBETPP

statements

Volume 2 Number 2 February 2001

A Brief History of Attempts to Monitor Testing

George F. Madaus
National Board on Educational Testing and Public Policy
Carolyn A. and Peter S. Lynch School of Education
Boston College

The idea of establishing standards for psychological testing or somehow monitoring the use of tests has a long history. As far back as 1895, the American Psychological Association (APA) appointed a committee to investigate the feasibility of standardizing mental and physical tests. During the 1900s, some psychologists prescribed specific standards for tests – for example, in 1924 Truman Kelley wrote that a test needed a reliability of 0.94 to be useful in evaluating individual accomplishment.[1] But organized efforts to standardize tests bore little fruit until mid-century.

Since then, such standards have proliferated. Notable among them is a series jointly sponsored by the APA, the American Educational Research Association (AERA), and the National Council on Measurements in Education (NCME). This series began in 1954 when the APA produced *Technical Recommendations for Psychological Tests and Diagnostic Techniques.*[2]

TM033207

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
M. Clarke
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

BEST COPY AVAILABLE

The National Board on Educational Testing and Public Policy, located in the Lynch School of Education at Boston College, is an independent body created to monitor testing in American education. The National Board provides research-based information for policy decision making, with special attention to groups historically underserved by the educational system. In particular, the National Board

- Monitors testing programs, policies, and products
- Evaluates the benefits and costs of specific testing policies
- Evaluates the extent to which professional standards for test development and use are met in specific contexts

This paper traces the history and development of the idea that testing and testing programs need the oversight of a monitoring body.

George Madaus is a Senior Fellow with the National Board on Educational Testing and Public Policy and the Boisi Professor of Education and Public Policy in the Lynch School of Education at Boston College.

Statements Series Editor: Marguerite Clarke

The AERA and the National Council on Measurements Used in Education (the forerunner of NCME) collaborated to produce the 1955 *Technical Recommendations for Achievement Tests*.[3] In 1966, and again in 1974 and 1985, the APA, AERA, and NCME issued revised versions of the technical recommendations called the *Standards for Educational and Psychological Testing* (the *Standards*).[4] In 1992 these three organizations began another revision of the Standards and a new version was released in 1999.[5]

We will not trace the evolution of these professional standards and ethical codes here. Instead, we focus on efforts to organize a means of monitoring testing. We describe proposals to that end in chronological order, from the first proposal for independent monitoring of tests in the 1920s, to similar proposals in the 1990s.

The proposals described include:

- Giles Ruch's proposal for a consumers' research bureau on tests;
- Oscar K. Buros' reviews of tests and efforts to establish a more active test monitoring agency;
- The APA's call for a Bureau of Test Standards and a Seal of Approval;
- The Project on the Classification of Exceptional Children's recommendation for a National Bureau of Standards for Psychological Tests and Testing; and
- The efforts of various organizations to establish standards for test development and use (e.g., the AERA, APA, and NCME *Standards for Educational and Psychological Testing* and the APA *Guidelines for Computer-based Tests and Interpretations*).

Ruch Proposal for Consumers' Research Bureau on Tests

As far as we know, the first call for an independent monitoring agency for testing came in 1925, from Giles M. Ruch. Ruch, a well-known author of numerous standardized tests, was concerned by the lack of information that test publishers provided and argued that "the test buyer is surely entitled to the same protection as the buyer of food products, namely, the true ingredients printed on the outside of each package."[7] Eight years later, Ruch had seen little improvement in the situation and proposed an external agency to evaluate tests:

> There is urgent need for a fact-finding organization which will undertake impartial, experimental, and statistical evaluations of tests – validity, reliability, legitimate uses, accuracy of norms, and the like. This might lead to the listing of satisfactory tests in the various subject matter divisions in much the same way that Consumers' Research, Inc. is attempting to furnish reliable information to the average buyer.[8]

Ruch's efforts to initiate such an organization were without success.

"The test buyer is surely entitled to the same protection as the buyer of food products, namely, the true ingredients printed on the outside of each package."

Buros' Reviews of Tests

The second, and much more successful, effort to monitor testing was begun by Oscar K. Buros in the 1930s. For over forty years, until his death in 1978, Buros directed the Buros Institute of Mental Measurements and through it a crusade to improve the quality of tests and their use. His wife, Luella, who assisted him, was instrumental in having the institute relocated to the University of Nebraska, where its work continues via publication of the *Mental Measurements Yearbook (MMY)*[9] and *Tests in Print (TIP)*[9] series.

Buros is known as the pre-eminent bibliographer of tests, and the publications he initiated have become the standard reference sources on tests.[10] Initially, however, he sought more active monitoring of testing. In the 1930s Buros echoed Ruch's call for a monitoring agency. He believed that neither commercial test publishers nor non-profit organizations such as the Cooperative Test Service and the sponsors of state testing programs could be unbiased critics of their own tests. He reported that he tried without success to start a test consumers' research organization.

BEST COPY AVAILABLE

The Buros Institute's . . . goal is to help test users by influencing test authors and publishers to produce better tests and to provide better information with them.

Buros then initiated the test review project that led to the *MMY*. When the first yearbook was published by Rutgers University in 1938, Buros still hoped for an external test monitoring agency. Clarence Partch, Dean of the School of Education, noted in his foreword that the School of Education hoped to establish a Test Users' Research Institute to evaluate tests and testing programs and serve as a clearinghouse for information on testing. This never came to pass.

The Buros Institute's work came to comprise the *MMY* and *TIP* series, and a series of monographs on tests in particular subject areas. The Institute also now maintains an on-line database with monthly updates of the publications. Its goal is to help test users by influencing test authors and publishers to produce better tests and to provide better information with them. This goal has remained essentially unchanged since 1938:

> Test authors and publishers will be impelled to construct fewer and better tests and to furnish a great deal more information concerning the construction, validation, use, and limitations of their tests.... Test users will be aided in setting up evaluation programs that will recognize the limitations and dangers associated with testing – and the lack of testing – as well as the possibilities.[1]

To that end, the Institute provides a list of available tests, information about them, critical reviews by independent persons from psychology, testing and measurement, and related fields, and bibliographies. The centerpiece of the Institute's work is the *MMY* series, of which the thirteenth and most recent yearbook was published in 1998. Each yearbook supplements the previous editions; it does not repeat information for tests previously reviewed that were not substantially revised in the interim. The *TIP* series is more bibliographical; each volume supersedes the previous one and lists all tests available for use with English-speaking subjects. The series also provides a master index to the Yearbooks. The most recent volume was published in 1999.

The monographs series reprints information from the *MMY*s and *TIP*s for particular types of tests. It has covered, for example, reading tests, personality tests, intelligence tests, social studies tests, and science tests.

BEST COPY AVAILABLE

The Buros Institute's work has been extremely successful in several respects. Its mere longevity is evidence of success; for most of its sixty-year history, it has supported itself by the sale of its publications. These are comprehensive and have a well-deserved reputation for objectivity based on the integrity of the editors and the independence of the reviewers.

The Institute's success story is tempered, however, by its failures and limitations. Its success at supporting itself was due to its nearly complete failure to find outside funding. Buros had some initial support, but this dried up early. By 1972, eight of the Institute's ten publications to that point had been published by the Gryphon Press, which consisted of Buros and his wife. Since Buros's death and the relocation of the Institute to the University of Nebraska-Lincoln, the publishing effort is apparently on a sounder basis, since the series is now distributed by the University of Nebraska Press.

Buros himself considered his life's work less than a complete success. In addition to the bibliographical and review functions of the Institute, Buros had pursued five objectives of a "crusading nature:"

- to impel test authors and publishers to publish better tests and to provide detailed information on test validity and limitations;
- to make test users aware of the value and limitations of standardized tests;
- to stimulate reviewers to think through more carefully their own beliefs and values relevant to testing;
- to suggest to test users better methods of appraising tests in light of their needs; and
- to urge suspicion of all tests unaccompanied by detailed data on their construction, validity, uses, and limitations.[17]

Buros called the results of these endeavors modest. He found that test publishers continued to market tests that failed to meet the standards of *MMY* and journal reviewers, and that at least half of them should never have been published. Exaggerated, false, or unsubstantiated claims were the rule. While test users were becoming somewhat more discriminating, a test – no matter how poor – that was nicely packaged and promised to do all sorts of things no test can do still found many gullible buyers.

While test users were becoming somewhat more discriminating, a test – no matter how poor – that was nicely packaged and promised to do all sorts of things no test can do still found many gullible buyers.

. . . the effects of tests cannot be divorced from the effects of testing. Indeed, some of the most serious problems of testing clearly have arisen not from shortcomings of the tests themselves, but rather from misuse of technically adequate products.

Failures aside, the Institute's work also has two major shortcomings. First, the critical reviews that are the core of the effort are produced by many people whose views on test quality inevitably vary. The editors of the eleventh *MMY* point out that readers should critically evaluate reviewers' comments on the tests since, while the reviewers are outstanding professionals in their fields, their reviews inevitably reflect their personal learning histories.

Second, the Buros publications have focused largely on tests and not on testing. They deal with the quality of the tests produced; but the effects of tests cannot be divorced from the effects of testing. Indeed, some of the most serious problems of testing clearly have arisen not from shortcomings of the tests themselves, but rather from misuse of technically adequate products.

Two Calls for a Bureau of Test Standards

There have been at least two other calls for a "bureau of test standards." The first came more than forty years ago from a committee of the APA. The APA is well known today for its part in creating the *Standards for Educational and Psychological Testing*. Less well known is that when it formed its original Committee on Test Standards in 1950, it also considered establishing a Bureau of Test Standards and a Seal of Approval. The Committee would have enforced its standards through the Bureau and by granting the Seal of Approval. The Committee was in fact established (and is now known as the APA Committee on Psychological Tests and Assessments, or CPTA), but the proposal for a Bureau and Seal apparently went nowhere. The records of the APA note simply that "the Council voted to take no action on these two recommendations, in view of the complicated problems they present."[13]

A quarter-century later a national commission recommended a similar body, but this time as a federal agency. Under the auspices of what was then the Department of Health, Education, and Welfare, the Project on the Classification of Exceptional Children was charged with examining the classifying and labeling of children who were handicapped, disadvantaged, or delinquent. The project report allowed that well-designed standardized tests could have value when used appropriately by skilled persons, but found that tests were too often of poor quality and misused, and that the "admirable efforts" of professional organizations and reputable test publishers did not "prevent widespread abuse."[14] The report stated:

> Because psychological tests... saturate our society and because their use can result in the irreversible deprivation of opportunity to many children, especially those already burdened by poverty and prejudice, we recommend that there be established a National Bureau of Standards for Psychological Tests and Testing.[15]

It further suggested that poor tests or testing could be as injurious to opportunity as impure food or drugs are injurious to health. The proposed Bureau would have set standards for tests, tests uses, and test users, acted on complaints, operated a research program, and disseminated its findings.

... poor tests or testing could be as injurious to opportunity as impure food or drugs are injurious to health.

BEST COPY AVAILABLE

What happened to the recommendation of this report? Apparently nothing. Edward Zigler, then Director of the Office of Child Development, who proposed the project, recalls only that "the recommendation... was never followed up."[16]

. . . it is evident that while ethical standards directly relevant to testing have diminished in number, technical standards have multiplied.

Joint Standards for Educational and Psychological Tests

In comparing the evolution of the APA ethical standards and the joint AERA-APA-NCME test standards (i.e., the *Standards*) from the 1950s through the mid-1980s, it is evident that while ethical standards directly relevant to testing have diminished in number, technical standards have multiplied. Some test publishers clearly have been paying heed to the joint test standards. For example, the Educational Testing Service (ETS) *Standards for Quality and Fairness*,[17] adopted by the ETS Trustees in the mid-1980s, reflect and adopt the *Standards*. Adherence to the *Standards for Quality and Fairness* is assessed through audit and subsequent management review and monitored by a Visiting Committee of persons outside ETS that includes educational leaders, testing experts, and representatives of organizations that have been critical of ETS.

But numerous small publishers violate the *Standards* (e.g., with regard to documenting validity and distributing test materials). Moreover, the connection between the *Standards* and test use is quite weak:

> There is much evidence that the test standards [i.e., the *Standards*] have limited direct impact on test developers' and publishers' practices and even less on test use.... [Yet]...there seems to be little professional enthusiasm for concrete proposals to enforce standards.... Professionals seem reluctant to set up regular...mechanisms for the enforcement of their standards in part because the notion of self-governance and professional judgment is part of [their] self-image.... As Arlene Kaplan Daniels has observed, professional "codes...are part of the ideology, designed for public relations and justification for the status and prestige which professions assume...."[18]

BEST COPY AVAILABLE

These conclusions seem to us still relevant to efforts since the mid-1980s to develop standards for testing. To illustrate this point, we cite two examples relating to "standards" promulgated for computerized testing, and for honesty or integrity testing. Before describing these two cases, we note that since the mid-1980s there have been several other initiatives to set standards for testing:

- In 1987 the Society for Organizational and Industrial Psychology developed the *Principles for Validation and Use of Personnel Selection Procedures.*[19]
- In 1988 the *Code of Fair Testing Practices* was completed.[20] It was developed by the Joint Committee on Testing Practices, initiated by AERA, APA, and NCME, but with members from other professional organizations. The *Code* was intended to be consistent with the 1985 *Standards;* it is limited to educational tests and was to be understandable by the general public.[21] It has been endorsed by numerous test publishers.
- In 1990 the American Federation of Teachers issued its *Standards for Teacher Competence in Educational Assessment of Students.*[22]
- In 1991 a National Forum on Assessment developed *Criteria for Evaluating Student Assessment Systems,* which was endorsed by more than five dozen national and regional education and civil-rights organizations. Subsequently, FairTest, one of the members of the National Forum, proposed requiring an Educational Impact Statement (similar to Environmental Impact Statements) before adoption of any new national testing system.[23]

. . . one of the members of the National Forum, proposed requiring an Educational Impact Statement (similar to Environmental Impact Statements) before adoption of any new national testing system.

Guidelines for Computer-based Tests and Interpretations

Because computerized tests and test interpretation were increasing rapidly in the 1980s, the APA decided to develop the 1986 *APA Guidelines for Computer-based Tests and Interpretations* (the *Guidelines*).[28] The *Guidelines* aimed to interpret the 1985 *Standards* as they relate to computer-based testing and interpretation, and to outline professional responsibilities in this field. They clearly specify that like paper-and-pencil tests, computer-based testing should undergo scholarly peer review. Guideline 31 states:

> Adequate information about the [computer] system and reasonable access to the system for evaluating responses should be provided to qualified professionals engaged in a scholarly review of the interpretive service. When it is deemed necessary to provide trade secrets, a written agreement of nondisclosure should be made.

The *Guidelines* aimed to interpret the 1985 *Standards* as they relate to computer-based testing and interpretation, and to outline professional responsibilities in this field. They clearly specify that like paper-and-pencil tests, computer-based testing should undergo scholarly peer review.

However, this guideline has had little effect on computerized testing, as noted in the introduction to the eleventh *MMY*:

> There has been a dramatic increase in the number and type of computer-based-test-interpretative systems (CBTI). We had considered publishing a separate volume to track the quality of such systems [but]... were frustrated...by the difficulty we encountered in accessing from the publishers the test programs and more importantly the algorithms in use by the computer-based systems.[29]

If even the Buros Institute, the pre-eminent agency for scholarly review of tests, has no access to computerized testing systems for review purposes, clearly the producers of these systems are not following the *Guidelines*.

BEST COPY AVAILABLE

Model Guidelines for Pre-employment Integrity Testing Programs

Another set of testing standards issued since the 1985 *Standards* is the *Model Guidelines for Pre-employment Integrity Testing Programs* (the *Model Guidelines*), developed by the Association of Personnel Test Publishers (APTP).[26] This association is a group of trade organizations that publish personnel tests. Most of the task force that developed the *Model Guidelines* is affiliated with personnel testing companies.

Two things are striking about these guidelines. First, while they refer to more widely recognized standards for testing (such as the 1985 *Standards*), they clearly have a promotional aura about them. For example, an introductory table, listing the "convenience issues," "main problems," and "main advantages" of various screening methods available to business and industry, clearly indicates that integrity tests are the best.

Second, these guidelines were developed on the heels of a marked increase in the sales of so-called honesty or integrity tests. In 1988 the U.S. Congress barred the use of polygraph tests to screen applicants for most jobs. Immediately thereafter, there was a flurry of advertising for paper-and-pencil honesty tests, which came to be quite widely used in some businesses. A 1990 survey showed, for example, that 30 percent of wholesale and retail trade businesses used such tests.[27]

At the same time, there was widespread concern about the validity and use of these tests. As a result, two investigations were launched in the late 1980s, one by the APA and one by the Office of Technology Assessment. Both turned out to be fairly critical of honesty testing (the OTA 1990 report more so than the APA Task Force 1991 study);[28] but oddly, the *Model Guidelines* make no reference whatsoever to either investigation. This omission can hardly be attributed to ignorance since many of the companies with which APA Task Force members are affiliated were surveyed in both studies.

Thus, although the *Model Guidelines* do contain some useful advice for potential developers and users of honesty or integrity tests, they are not an independent or scholarly effort. Indeed, one observer has suggested that the APTA *Model Guidelines* might be viewed as an attempt by a trade organization not just to improve the practices of personnel test publishers, but also to help fend off more active and independent monitoring of this segment of the testing marketplace.

The concept of monitoring tests and the impact of testing programs on individuals and institutions . . . was not translated into practice until the formation in 1998 of the National Board on Educational Testing and Public Policy.

Conclusion

The concept of monitoring tests and the impact of testing programs on individuals and institutions has a long history. Its merit is commonly acknowledged. Nevertheless, it was not translated into practice until the formation in 1998 of the National Board on Educational Testing and Public Policy. The National Board, funded by a startup grant from the Ford Foundation, has finally begun the process of independently monitoring tests and testing programs that has been called for since the 1920s.

notes

1. Kelley, T. L. (1924). *Statistical method*. New York: MacMillan.
2. American Psychological Association. (1954). *Technical recommendations for psychological tests and diagnostic techniques*. Washington, DC: American Psychological Association.
3. American Educational Research Association, Committee on Test Standards, and National Council on Measurements Used in Education. (1955). *Technical recommendations for achievement tests*. Washington, DC: American Educational Research Association.
4. American Educational Research Association, American Psychological Association, and the National Council on Measurement in Education. (1985). *Standards for educational and psychological testing*. Washington, DC: American Psychological Association.
5. See: American Educational Research Association, American Psychological Association, and the National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, DC: American Psychological Association. The three organizations that have sponsored the standards also have developed their own ethical codes. For example, the APA issued ethical standards or principles (and most recently a code of conduct) in 1953, 1958, 1963, 1968, 1977, 1979, 1981, 1990, and 1992. Also, the AERA issued its first set of ethical standards in 1992, and the NCME developed *The Code of Professional Responsibility in Educational Measurement* in 1995 (see Schmeiser, C. B. (1992). Ethical codes in the professions. *Educational Measurement: Issues and Practice*, 11(3), 5-11).
6. Ruch, G. M. (1925). Minimum essentials in reporting data on standard tests. *Journal of Educational Research*, 12, 349-358.
7. Ibid.
8. See: Buros, O. K. (Ed.). (1938, 1940, 1949, 1953, 1959, 1965, 1974, 1978); Mitchell, J. V. (Ed.). (1985); Conoley, J. C., and Kramer, J. J. (Eds.). (1989, 1992); Conoley, J.C., and Impara, J.C., (Eds.). (1995); Impara, J.C., and Plake, B.S. (Eds.). (1998). *Mental measurement yearbook*. Highland Park, NJ: Gryphon Press.
9. See: Buros, O.K. (Ed.). (1961, 1974); Mitchell, J. (Ed.). (1983); Murphy, L., Close Conoley, J., and Impara, J. (Eds.). (1994); Murphy, L., Impara, J., and Plake, B. (Eds.). (1999). *Tests in print*, 1-5. Highland Park, NJ: Gryphon Press.
10. In the 1980s an alternative compendium of reviews of tests, called *Test Critiques*, was begun by the Test Corporation of America, a subsidiary of Westport Publishers of Kansas City, MO. This series has not attained nearly the stature of the Buros series for at least four reasons. First, it is of much more recent vintage. Second, it provides reviews of only better known and widely used tests. Third, it contains only reviews and not the extensive bibliography of the Buros series. Fourth, as the product of a commercial publishing house, it seems unlikely to attain a reputation like that of the Buros series for independent scholarship. This title currently consists of Volumes I-X (http://www.siu.edu/colleges/AS/PSY/Tests1.html). See: Keyser, D. J., and Sweetland, R. C. (Eds.). (1984-current). *Test critiques*. Kansas City, MO: Test Corporation of America.
11. P. XI. Buros, O. K., Ed. (1972). *The seventh mental measurements yearbook*. Highland Park, NJ: Gryphon Press.
12. Ibid, p. XXVII.
13. P. 546, emphasis added. Adkins, D. C. (1950). Proceedings of the Fifty-eighth Annual Business Meeting of the American Psychological Association, Inc., State College, Pennsylvania. *The American Psychologist*, 5, 544-575.
14. P. 237. Hobbs, N. (1975). *The futures of children*. San Francisco: Jossey-Bass.
15. Ibid.
16. Zigler, E. (1991). Letter to Kenneth B. Newton.
17. Educational Testing Service. (1987). *ETS standards for quality and fairness*. Princeton, NJ: Educational Testing Service.
18. P. 49. Daniels, A. K. (1973). How free should professionals be? *The professions and their prospects* (pp.39-57). Beverly Hills, CA: Sage.
19. Society for Organizational and Industrial Psychology, Inc. (1987). *Principles for the validation and use of personnel selection procedures*. College Park, MD: Society for Organizational and Industrial Psychology, Inc.
20. American Educational Research Association, American Psychological Association, and the National Council on Measurement in Education Joint Committee on Testing Practices. (1988). *Code of fair testing practices*. Washington, DC: American Psychological Association.

BEST COPY AVAILABLE

21 Fremer, J. E., Diamond, E., and Camara, W. (1989). Developing a code of fair testing practices in education. *American Psychologist*, 44(7), 1062-1067.

22 American Federation of Teachers, National Council on Measurement in Education, and National Education Association. (1990). *Standards for teacher competence in educational assessment of students*. Washington, DC: American Federation of Teachers.

23 Neill, M. (September 23, 1992). Assessment and the educational impact statement. *Education Week*, 12(3).

24 American Psychological Association. (1986). *Guidelines for computer-based tests and interpretations*. Washington, DC: The American Psychological Association.

25 P. Xi. Conoley, J. C., and Kramer, J. J. (Eds.). (1992). *The eleventh mental measurements yearbook*. Lincoln, NE: Buros Institute of Mental Measurements (distributed by the University of Nebraska Press).

26 Association of Personnel Test Publishers. (1990). *Model guidelines for preemployment integrity testing programs*. Washington, DC: Association of Personnel Test Publishers.

27 American Management Association. (1990). *The AMA 1990 survey on workplace testing*. New York: American Management Association.

28 Office of Technology Assessment. (1990). *The use of integrity tests for pre-employment screening*. Washington, DC: Office of Technology Assessment.

American Psychological Task Force on the Prediction of Dishonesty and Theft in Employment Settings. (1991). *Questionnaires used in the prediction of trustworthiness in pre-employment selection decisions: An A.P.A. task force report*. Washington, DC: American Psychological Association.

BEST COPY AVAILABLE

BEST COPY AVAILABLE

Other National Board on Educational Testing and Public Policy Publications

statements

An Agenda for Research on Educational Testing.
Marguerite Clarke, George Madaus, Joseph Pedulla, and Arnold Shore

The Gap between Testing and Technology in Schools.
Michael Russell and Walter Haney

High Stakes Testing and High School Completion.
Marguerite Clarke, Walter Haney, and George Madaus

Guidelines for Policy Research on Educational Testing.
Arnold Shore, George Madaus, and Marguerite Clarke

The Impact of Score Differences on the Admission of Minority Students: An Illustration.
Daniel Koretz

Educational Testing as a Technology.
George F. Madaus

web statements

Mode of Administration Effects on MCAS Composition Performance for Grades Four, Eight, and Ten.
Michael Russell and Tom Plati

monographs

Cut Scores: Results May Vary.
Catherine Horn, Miguel Ramos, Irwin Blumer, and George Madaus

NBETPP
Visit us on the World Wide Web at nbetpp.bc.edu for more articles, the latest educational testing news, and information on NBETPP.

About the National Board on Educational Testing and Public Policy

Created as an independent monitoring system for assessment in America, the National Board on Educational Testing and Public Policy is located in the Carolyn A. and Peter S. Lynch School of Education at Boston College. The National Board provides research-based test information for policy decision making, with special attention to groups historically underserved by the educational systems of our country. Specifically, the National Board

- Monitors testing programs, policies, and products
- Evaluates the benefits and costs of testing programs in operation
- Assesses the extent to which professional standards for test development and use are met in practice

This National Board publication series is supported by a grant from the Ford Foundation.

The National Board on Educational Testing and Public Policy
Lynch School of Education, Boston College
Chestnut Hill, MA 02467

Telephone: (617)552-4521 • Fax: (617)552-8419
Email: nbetpp@bc.edu

Visit our website at **nbetpp.bc.edu** for more articles, the latest educational news, and information on NBETPP.

NBETPP

The Board of Overseers

Antonia Hernandez
President and General Council
Mexican American Legal Defense and Educational Fund

Harold Howe II
Former U.S. Commissioner of Education

Paul LeMahieu
Superintendent of Education
State of Hawaii

Peter Lynch
Vice Chairman
Fidelity Management and Research

Faith Smith
President
Native American Educational Services

Gail Snowden
Managing Director, Community Banking
FleetBoston Financial

Peter Stanley
President
Pomona College

Donald Stewart
President and CEO
The Chicago Community Trust

BOSTON COLLEGE